Structure of the MeCP2-TBLR1 complex reveals a molecular basis for Rett syndrome and related disorders

Structure of the MeCP2-TBLR1 complex reveals a novel molecular basis for Rett syndrome

Valdeko Kruusvee*[1], Matthew J. Lyst*[1], Ceitidh Taylor[1], Žygimanté Tarnauskaité[2], Adrian P. Bird[§1] and Atlanta G. Cook[§1]

1. Wellcome Trust Centre for Cell Biology, University of Edinburgh, Max Born Crescent, Edinburgh, EH9 3BF, United Kingdom

2. MRC Human Genetics Unit MRC IGMM, University of Edinburgh, Western General Hospital, Crewe Road, Edinburgh EH4 2XU, United Kingdom

*These authors contributed equally to the work

[§]Address correspondence to: A.Bird@ed.ac.uk or atlanta.cook@ed.ac.uk

Wellcome Trust Centre for Cell Biology,

University of Edinburgh,

Max Born Crescent,

Edinburgh,

EH9 3BF,

United Kingdom

Classification: BIOLOGICAL SCIENCES: Biochemistry

Running title: Structure of the MeCP2-TBLR1 complex

Abstract

Rett syndrome (RTT) is an X-linked neurological disorder caused by mutations in the *MECP2* gene. The majority of RTT missense mutations disrupt the interaction of MeCP2 with DNA or the NCoR/SMRT co-repressor complex. Here we show that the "NCoR/SMRT Interaction Domain" (NID) of MeCP2 directly contacts TBL1 and TBLR1, two paralogues that are core components of NCoR/SMRT. We determine the co-crystal structure of the MeCP2 NID in complex with the WD40 domain of TBLR1 and confirm by *in vitro* and *ex vivo* assays that mutation of interacting residues of TBLR1 and TBL1 disrupts binding to MeCP2. Strikingly, the four MeCP2 NID residues mutated in RTT are those that make most extensive contacts with TBLR1. Moreover, missense mutations in the gene for TBLR1 that are associated with intellectual disability also prevent MeCP2 binding. Our study therefore reveals the molecular basis of an interaction that is crucial for optimal brain function.

Significance statement

The nuclear protein MeCP2 links epigenetics, brain function and neurodevelopmental disease. Mutations in the *MECP2* gene cause Rett syndrome, making it imperative to determine its mechanism of action. One protein domain targets MeCP2 to methylated DNA, but little was known about a second essential domain except that it interacts with a gene silencing complex. We determined that TBL subunits of NCoR/SMRT bind MeCP2 and solved the structure of the binary complex. Strikingly, amino acids mutated in Rett syndrome are precisely those that intimately contact the TBL subunits. Furthermore, mutations in TBL proteins that cause intellectual disability prevent the interaction with MeCP2. This suggests that the TBL-MeCP2 interaction is essential for brain function and may be accessible to therapeutic modulation.

Introduction

The methyl-CpG binding protein MeCP2 is a chromatin-associated factor that is highly expressed in the brain (1, 2). Loss-of-function mutations in MeCP2 lead to the severe pediatric neurological disorder, Rett syndrome (RTT) (3), which affects around 1 in 10,000 girls. In addition, extra copies of the gene cause MeCP2 duplication syndrome (4, 5), a distinct intellectual disability disorder that predominantly affects males (6). The importance of MeCP2 for brain function has prompted investigation of its molecular function. The protein was identified because of its ability to bind DNA via its methyl-CpG binding domain (MBD) (1, 7) and many residues mutated in RTT disrupt this interaction (8-11). In addition to DNA, numerous protein partners of MeCP2 have been reported (11). Of these, only the interactions with the nuclear receptor co-repressor complexes NCoR and its close relative SMRT (silencing mediator of retinoic acid and thyroid receptors) (NCoR/SMRT) (12-14) are known to be disrupted by RTT missense mutations (13). Accordingly, Rett missense mutations outside the MBD are found clustered within the so-called NCoR/SMRT Interaction Domain (NID, Fig. 1a, S1a) (13). One NID mutation, R306C, causes a severe RTT-like phenotype in mice (13, 15, 16). Moreover, mouse models of MeCP2 duplication syndrome suggest that both the MBD and the NID must be intact for this adverse molecular pathology to develop (16, 17).

These findings indicate that the NID mediates an essential function of MeCP2. This function may be recruitment of the NCoR/SMRT co-repressor complexes to chromatin. Alternatively, the NID may perform other essential roles, such as DNA binding, whose loss leads to RTT(16). In an attempt to distinguish these possibilities we investigated the molecular basis of the interaction between MeCP2 and NCoR/SMRT. The latter are ~1.2 MDa multi-subunit complexes that include NCoR1 (and/or SMRT), HDAC3, GPS2 and TBL1 (and/or its paralogue TBLR1) (18-21). We show here that the TBL1 and TBLR1 subunits (89% identical in amino acid sequence in humans; Fig. S1b) are direct MeCP2 binding partners. The crystal structure of a minimal MeCP2-TBLR1 complex at 2.5 Å resolution reveals that the four MeCP2 NID residues most

intimately contacting TBLR1 are the same residues that are mutated in RTT. This finding strongly suggests that sustaining this interaction is the primary function of the NID. Mutations in the TBLR1 WD40 domain that disrupt MeCP2 NID binding *in vitro* also prevent recruitment of TBL1 to heterochromatic foci in cells. Finally, we show that intellectual disability-associated missense mutations in TBLR1 disrupt its interaction with MeCP2, supporting the view that this interaction is critical for brain function.

Results

MeCP2 interacts with NCoR/SMRT and TBL1 in cells

Previously, we showed that an MeCP2-derived NID peptide could pull-down an NCoR1 fragment spanning residues 197 to 598 (13). To define the MeCP2 binding site on the NCoR/SMRT complex, we used a similar strategy with FLAG-tagged NCoR1 fragments expressed in HeLa cells (Fig. 1a,b). Like NCoR1(197-598), NCoR1 fragments spanning residues 197-490 and 227-490 were bound efficiently by a wild-type NID peptide corresponding to residues 285-313 (MeCP2-NID, Fig. 1a,b). Two further N-terminal truncations showed reduced binding (residues 241-490) or failed to bind (297-490) (Fig. 1b). None of these constructs bound to a control NID peptide containing the RTT mutation K305R, confirming specificity. We noted that residues 227-297 of NCoR1, which appear to be involved in binding MeCP2, overlap with the NCoR1 region that interacts with TBL1 and TBLR1 (NCoR1 residues 247-256) (22). This suggested that the MeCP2-NCoR1 interaction might be indirect, mediated by interaction of MeCP2 with endogenous TBL1/TBLR1 present in cell lysates. We used a co-immunoprecipitation (co-IP) assay to test whether the ability of FLAG-tagged NCoR1 deletion fragments to bind to endogenous TBLR1 matched their ability to bind MeCP2. Indeed the NCoR1(241-490) and (297-490) fragments showed reduced and absent binding to TBLR1, respectively (Fig. 1c). The correlation between pull-down efficiency of NCoR1 fragments by the NID peptide and TBLR1 suggests that the MeCP2-NCoR1 interaction is bridged by TBL1/TBLR1.

MeCP2 binds directly to TBL1 and TBLR1

TBLR1 and TBL1 have an N-terminal tetramerisation domain, which interacts with NCoR1, linked to a C-terminal WD40 domain (Figs. 1a,2a) (27). To determine the minimal fragment of TBL1 that binds MeCP2, we expressed FLAG-tagged full-length, N-terminal and C-terminal fragments of TBL1 in HeLa cells (Fig. 2a,b). Full-length TBL1 and a C-terminal fragment of TBL1 (residues 176-527) were efficiently precipitated by the MeCP2 NID peptide, whereas the N-terminal domain of TBL1 (residues 1-181) was not (Fig. 2b). None of the deletion mutants bound to a peptide carrying the K305R RTT mutation. Furthermore, shortening of the WD40 domain to disrupt the first or last repeat (FLAG-TBL1 (200-527) or FLAG-TBL1 (176-488)) disrupted the interaction with NID peptide (Fig 2b). These deletions likely destabilize the WD40 domain, indicating that an intact TBL1 WD40 domain is required to bind MeCP2.

We next asked whether the NID peptide could interact directly with purified recombinant mouse TBL1 and TBLR1 C-terminal WD40 domains (CTDs Fig. 2c, Fig. S2a). TBL1 WD40 domain was efficiently pulled down by wild-type NID but not by RTT mutant peptides or a peptide containing phosphoThr308, a modification known to interfere with NID binding (Fig. 2c)(23). An MeCP2 peptide up to and including residue 308, corresponding to a hypomorphic mutation (24) with reduced NCoR binding (13), also showed weakened binding to TBL1 (Fig. 2c). These results confirm that MeCP2 NID can bind directly to the WD40 domain of TBL1, both in cell extracts and *in vitro*. Using surface plasmon resonance (SPR) we measured binding constants (K_D) for alternative NID peptides (NID_s, residues 285-309; or NID residues 285-313) with the mouse TBLR1 carboxy-terminal domain (TBLR1-CTD; residues 134-514) (Fig. 2d, Fig. S2b). Wild-type TBLR1-CTD bound MeCP2-NID and MeCP2-NID_S with apparent K_Ds of 9.5 ± 0.5 μM and 12.9 ± 0.8 μM, respectively. Similar binding constants suggest that MeCP2 residues up to and including 309 are sufficient for binding. Finally, we find that a minimal 11-mer peptide, spanning residues 298 to 309 of MeCP2, is sufficient to bind TBLR1 in fluorescence anisotropy assays (Fig. S2c). While the binding affinity of the binary interaction is relatively low, binding *in vivo* might be enhanced by the tetrameric structure of TBL1/TBLR1.

TBL1 chimeras show differential binding to MeCP2

As progressive deletions that truncate the WD40 β-propeller domain destabilise the fold, we devised a "phylogenetic" strategy to identify residues required for MeCP2 binding. MeCP2 is only found in vertebrates, but TBL1 homologues are present throughout metazoans. By identifying evolutionarily remote TBL1 homologues, we were able to construct chimaeric TBL1 proteins that do not bind MeCP2, while retaining integrity of the WD40 domain. A *Drosophila* orthologue, Ebi (83% identity to human and mouse) was successfully pulled down by MeCP2 NID peptide, but a sponge ortholologue (*Amphimedon queenslandica*: 58% identity with human and mouse) was not recovered (Fig S3a). To map differences between mouse and sponge sequences, approximate quarters of the mouse TBL1 WD40 sequence were replaced with equivalent segments from sponge (Fig. 3a). Replacing the second quarter of the mouse sequence (TBL1-chimera1) or the first and last quarters (TBL1-chimera3) showed wild-type binding to MeCP2 peptide, while replacement of the third quarter (TBL1-chimera2) abolished binding (Fig. 3a). To refine the map, we generated a series of mouse TBL1 mutants where up to four consecutive residues were altered to equivalent residues in the sponge third quarter sequence (Fig. S3b,c). Only mutants with substituted 362-364 and 405-406 lost binding (Fig. S3c). Further mapping identified E364G as a single missense mutation in TBL1 that abolishes the MeCP2 interaction (Fig. 3b), but importantly did not prevent incorporation of TBL1 into the NCoR/SMRT complex (Fig. S3d). This residue maps to a surface of human TBLR1 (hTBLR1) that is associated with peptide ligand binding in related proteins (Fig. S3e) (25, 26).

MeCP2 NID binds the top surface of TBLR1 WD40 domain

To fully define the intermolecular interaction we determined a co-crystal structure of mouse TBLR1 C-terminal domain (TBLR1-CTD) with MeCP2 NID_S peptide. Crystals diffracted to 2.5 Å in space group *P*1. The structure was solved by molecular replacement using the coordinates of hTBLR1 as a search model (99.5% identity, PDB ID 4LG9, (26)). Four molecules of TBLR1-CTD are present in the asymmetric unit. After initial refinement, Fo-Fc maps showed clear electron density in two copies of TBLR1-CTD into which the NID_S peptide was built. The final model showed good stereochemistry (Table S1) (27). Chain A is the best-defined copy of TBLR1-CTD, where residues 155-513 were visible with the exception of short loops at residues 203-204 and 217-223. Chains C and D lacked peptide and displayed poor electron density. While the backbone of these chains could be discerned, many loops and side chains could not be modelled. We therefore restrict our discussion to TBLR1 chain A. Only residues 296 to 307 of the NID_S peptide could be modelled with confidence (Fig. S4a).

The TBLR1 WD40 domain is an eight-bladed β-propeller domain with a central channel (Fig. 4a,b) (26). As in other WD40 domains, the C-terminal β-propeller contains a β-strand from the first WD40 repeat sequence (Fig. 4a) (28). The NID_S peptide binds the top face of the central channel, a common interaction interface (25, 28). At the deepest point the peptide is approximately 10 Å inside the channel (Fig. 4b), burying a surface area of 532 $Å^2$ on TBLR1 (Fig. 4c). The TBLR1 interaction surface is highly electronegative (Fig. 4d), enabling recognition of the cluster of basic residues in the NID. Surface conservation calculations using vertebrate TBLR1 sequences show that the top face is conserved, consistent with a role in protein-protein interactions (Fig. 4e) (29).

Four key residues of MeCP2 interact with TBLR1

Side chains of residues 301-306 of MeCP2 were well resolved suggesting that this region forms the core of the interaction (Fig. 5a,b). The most striking feature of the NID is a tight backbone turn stabilised by $R306^{MeCP2}$ via a network of intra- and inter-molecular interactions within the complex. This residue forms a salt bridge with $E171^{TBLR1}$ and close contacts with the backbone carbonyl groups of $T299^{MeCP2}$ and $L301^{MeCP2}$ (Fig. 5a). These interactions and stereochemical constraints imposed by $P302^{MeCP2}$ enable the peptide to enter the central channel (Fig. 5a). $P302^{MeCP2}$ is part of a stretch of hydrophobic residues in the N-terminal part the MeCP2 core interacting region, which also includes $V300^{MeCP2}$, $L301^{MeCP2}$ and $I303^{MeCP2}$. $P302^{MeCP2}$ is buried in a hydrophobic pocket created by three aromatic side chains: $Y395^{TBLR1}$, $F420^{TBLR1}$ and $Y446^{TBLR1}$ (Fig. 5b). The hydroxyl group of $Y446^{TBLR1}$ provides a peptide backbone contact, forming a hydrogen bond with the main chain carbonyl of $V300^{MeCP2}$ (Fig. 5b).

In addition to $R306^{MeCP2}$, the C-terminal portion of the core NID has two further basic residues: $K304^{MeCP2}$ and $K305^{MeCP2}$. Both residues have extensive hydrogen bonding and ionic interactions with TBLR1. $K304^{MeCP2}$ and $K305^{MeCP2}$ form hydrogen bonds with main chain and side chain moieties of $N353^{TBLR1}$, respectively (Fig. 5b). $K305^{MeCP2}$ is further stabilised by

hydrogen bonding with D369^{TBLR1} and an ionic interaction with E351^{TBLR1}, which lies 3.1 Å away (Fig. 5b). Notably, E351^{TBLR1} is equivalent to E364 in TBL1, which proved to be essential for the interaction with MeCP2 based on our earlier phylogenetic mapping (Fig. 3b, S3e). This mode of recognition differs from other WD40 domain peptide interactions involving basic motifs, such as recognition of modified arginine and lysine residues in histone tails, where aromatic cages stabilize the positive charges (Fig. S4b) (30, 31).

RTT mutations disrupt individual MeCP2-TBLR1 contacts

Of the 25 MeCP2 NID_S peptide residues present in the co-crystal structure (Fig. 1a), only four make extensive interactions with TBLR1. Strikingly, missense substitutions at each of these four residues (P302, K304, K305 and R306) are reported in multiple independent cases of RTT. On the other hand, pathological substitutions at other NID residues are not reported in a database with over 1000 RTT mutations (32). The new structure explains in molecular detail why known RTT mutations in the NID likely prevent association with the NCoR/SMRT co-repressor complex.

The most common RTT mutation in this region, R306C^{MeCP2}, disrupts TBLR1 binding in two ways. Firstly, loss of the arginine side chain would prevent the formation of a salt bridge with E171^{TBLR1} (Fig. 5c). Secondly, the mutant cysteine side chain is short and lacks the hydrogen bonding capacity of the arginine guanidinium group, and so would be unable to stabilize the peptide conformation (Fig. 5c). RTT mutation K304E^{MeCP2} is a charge reversal mutation that likely prevents binding of the peptide by electrostatic repulsion (Fig. 5d). Two further RTT mutations, P302R^{MeCP2} and K305R^{MeCP2} introduce larger side chains than those present in the wild-type sequence. The arginine side chain of P302R^{MeCP2} could not be accommodated in the hydrophobic pocket where P302^{MeCP2} binds and would likely also disrupt the kinked backbone conformation of the peptide (Fig. 5e). K305R^{MeCP2} is superficially a conservative substitution. Given the precise arrangement of the MeCP2-NID_S peptide in the TBLR1 central channel, however, the longer arginine side chain creates a steric clash that would not be accommodated (Fig. 5f).

TBLR1 mutants disrupt interactions with MeCP2

To assess the importance of the TBLR1 residues that contact MeCP2, we tested a series of TBLR1 site-directed mutants using the SPR assay (Fig. 6a). E171A^{TBLR1} and E171Q^{TBLR1} assessed the importance of ionic interactions and hydrogen bonds with R306^{MeCP2} and showed a greater than four-fold reduction in binding to MeCP2 (Fig. 6b). These results suggest that the salt bridge with R306^{MeCP2} is more important than hydrogen bonding. Similarly, E351A^{TBLR1} showed weaker binding, whereas E351D^{TBLR1} retained wild-type affinity for MeCP2 (K_D = 11.4 ± 0.5 μM), indicating the importance of ionic interactions with K305^{MeCP2} (Table 1). D369A^{TBLR1}, which is also expected to prevent binding to K305^{MeCP2}, and D313N^{TBLR1}, which should affect the electronegative binding surface, both abolished binding to NID peptide (Table 1, Fig. 6b,S5a). Finally, Y446F^{TBLR1} mutation tested the importance of the main chain interactions with V300^{MeCP2}. The Y446F^{TBLR1} mutant showed close to wild-type binding, indicating a minor contribution to the interaction. As a

control $C214^{TBLR1}$, a residue distant from the MeCP2 binding site, was mutated to serine, and this mutant bound similarly to wild-type TBLR1. Only small variations in melting temperature of the mutant proteins were observed (deviations of ± 3ºC from wild-type Tm, 62ºC) in thermal denaturation assays (Fig. S5b), indicating that structural integrity of the proteins is unaffected. These results corroborate the structural data, as implicated residues are indeed required for efficient TBLR1 binding to MeCP2.

TBL1 mutants fail to co-localise with MeCP2

The structural analysis used only the TBLR1-CTD and a NID peptide. To validate these findings in living cells with full-length proteins, we tested for recruitment of TBL1/TBLR1 mutants to heterochromatin by MeCP2. We previously demonstrated that whereas TBL1-mCherry alone accumulates in the cytoplasm of NIH-3T3 cells, it is targeted to densely methylated heterochromatin when co-expressed with EGFP-MeCP2 (13). The interaction is dependent on the presence of a functional NID and is disrupted by the missense RTT mutation $R306C^{MeCP2}$ (13). As TBLR1-mCherry constructs gave a weak fluorescent signal compared with TBL1-mCherry, we introduced the equivalent mutations into TBL1 (Fig. 6a). Unlike wild-type TBL1-mCherry, TBL1 point mutants E184A, D326N, E364A and D382A showed little or no co-localisation with MeCP2 at heterochromatic foci (Fig. 6c). Mutations for which corresponding mutations in TBL1 do not affect TBLR1 binding *in vitro*, such as Y459F, showed a wild-type co-localisation (Fig. 6c). The correlation between the effect of mutations on protein distribution *ex vivo* and the effect on binding *in vitro* suggests that the molecular recognition observed with minimal interacting domains recapitulates the behaviour of full-length proteins in cells. Furthermore, all TBL1 mutants were able to co-IP HDAC3 when transfected into cells, further supporting the conclusion that these mutants retain a native fold (Fig. S5c).

Disease-associated TBLR1 mutations block MeCP2 binding

If a direct interaction between the MeCP2 and TBLR1 is critical for MeCP2 function, we expected that mutations in TBLR1 or TBL1 might also give rise to neurological disease. The database of genomic variation and phenotype in humans using Ensembl resources (DECIPHER (33)) revealed six missense mutations in *TBL1XR1*, the gene encoding TBLR1: H441R, H213Q, A311P, D369E, D370Y and P444R associated with intellectual disability and developmental delay. An additional missense mutation in *TBL1XR1*, L282P, was reported to be associated with sporadic autism spectrum disorder (34). All of these mutations map to the WD40 domain of TBLR1 (Fig. 7a, S6a). Five mutations (H213Q, D370Y, H441R, A311P and L282P) are likely to disturb the entire fold as they map either to the WD40 consensus sequence (25, 28) or the hydrophobic core of the protein (35).

Two mutations, P444R and D369E, were of more interest, as they map close to the MeCP2 binding site. Introduction of a large positively charged residue at P444 might destabilise the hydrophobic binding site formed by Y395, F420 and Y446 and thus interfere with MeCP2 binding (Fig. S6a), while the apparently conservative D369E mutation might block interaction with

K305^{MeCP2} (Fig. S6a). To test whether these two mutations prevent interaction with MeCP2, we co-precipitated wild-type or mutant FLAG-TBL1-CTD with MeCP2 NID peptide. We included E184A as a non-binding control. Neither D382E^{TBL1} (D369E^{TBLR1}) nor P457R^{TBL1} (P444R^{TBLR1}) is efficiently pulled down by MeCP2 NID peptide (Fig. 7b). Furthermore, these two mutants are not recruited to heterochromatic foci by MeCP2 in cells (Fig. 7c). Both of these mutants are incorporated efficiently into NCoR/SMRT complexes (Fig. S6b) and show only small effects on thermal stability (Fig. S6c), indicating that the mutations did not affect the overall structure of the protein. Consistent with the pull-down assay, the D369E mutation substantially reduced binding to the NID peptide while P444R had a milder effect (Fig. S6d). These clinically important mutations in TBL1/TBLR1 therefore interfere with binding to MeCP2 in cells.

Discussion

The clustering of RTT mutations indicates that there are two functionally important regions in MeCP2. While the role of the methylated DNA binding domain is well established, the effect of mutations that cluster in the NID was uncharacterised at the molecular level. Previous studies showed that the MeCP2 NID is required for interactions with the NCoR/SMRT complex in cells. Here we show that the NCoR/SMRT complex subunit TBL1, and its paralogue TBLR1, bind MeCP2 directly via a WD40 domain (Fig. 2). To our knowledge, this location is unique among known NCoR/SMRT recruiters, which consistently interact with the NCoR/SMRT scaffold proteins themselves.

The most striking aspect of the co-crystal structure is that all residues in the MeCP2 NID that are mutated in RTT play a direct role in binding TBLR1. The exact coincidence between the four residues crucial to the MeCP2–TBL1/TBLR1 interaction and four residues mutated in RTT makes it highly likely that TBL1 and TBLR1 are essential partners of MeCP2. It is notable that over 200 independent cases of the R306C mutation are reported, whereas mutation of P302 to R, L or A has been seen 23 times (32). Missense mutations at K305 and K304 are rarer, having been reported only 4 and 2 times respectively. Importantly, mutations at each of these amino acids are almost certainly causal due to their demonstrated absence in both parents of at least one patient (13, 32).

Given the importance of the MeCP2-TBL1/TBLR1 interaction, it might be expected that mutations in TBL1 or TBLR1 would lead to RTT. TBLR1 missense mutations have been reported in cases of developmental delay and sporadic autism spectrum disorder (33, 34), but so far not in patients diagnosed with RTT. Intriguingly, at least two of these mutations disrupt the interaction with MeCP2 in pull-down assays and in cells (Fig. 7). Differences in clinical presentation between *MECP2* and *TBL* mutations might be expected, as the *MECP2* gene is X-linked and therefore X chromosome inactivation in females leads to an interspersion of functionally wild-type and mutant cells. This kind of mosaicism does not apply to the autosomal gene *TBL1XR1* where mutations would affect all cells. It may also be relevant that TBL1 and TBLR1 exist in the same multi-protein complex, and so may partially complement each other's function. These considerations may lead to differences between phenotypic and/or clinical consequences of *MECP2* and

TBL mutations. An additional possibility is that TBL proteins interact with proteins other than MeCP2 and effects of mutations on these interactions could also contribute to neurological phenotypes in humans. Future work is needed to address these possibilities.

The combination of structural, biochemical and functional insights presented here provide strong evidence that the tripartite interaction between DNA, MeCP2 and TBL1/TBLR1 is crucial for brain function. Our co-crystal structure will inform future therapeutic strategies to combat RTT and other *MECP2*-related disorders. In particular, we note that small molecules that specifically interfere with the interaction may be useful in treating the severe neurological condition *MECP2* Duplication syndrome.

Materials and Methods

Co-immunoprecipitation and pull-down assays

Cells transiently transfected with p3xFLAG CMV 10 vectors encoding NCoR1 and TBL1 fragments were Dounce homogenized in NE10 buffer (20 mM HEPES.NaOH (pH 7.5), 10 mM NaCl, 1 mM $MgCl_2$, 0.1% Triton X-100 (v/v), protease inhibitors (Roche), 15 mM β-mercaptoethanol) then treated at 25 °C for 5 min with 250 units benzonase (Sigma) per 10^7 cells. The NaCl concentration was adjusted to 150 mM. After mixing for 20 min, lysates were cleared by centrifugation at 16,000 x g for 20 min. Pull-downs and co-immunoprecipitations were performed as described previously (13). Antibodies used were M2 anti-FLAG (Sigma) and anti-mCherry (Chromotek). For Western blots antibodies used were anti-HDAC3 (Santa Cruz, sc-11417), anti-TBLR1 (Bethyl, A300-408A), M2 anti-FLAG (Sigma) and anti-mCherry (Abcam ab167453).

Protein expression and purification

TBLR1-CTD (residues 134-514) was expressed as an N-terminal 6xHis tagged protein in Sf9 cells infected at a density of 2 x10^6 cells/ml with 1:50 virus:medium using a V2 stock. After three days, infected cells were resuspended in 20 mM Tris-HCl pH 7.5, 500 mM NaCl, 10 mM imidazole, 1 mM Pefabloc (Roche) supplemented with a complete EDTA-free protease inhibitor cocktail tablet (Roche). Cells were sonicated on ice and the lysate cleared by centrifugation at 50,000 x *g* at 4˚C for 45 min. His-tagged protein was batch purified using Ni^{2+}-NTA resin, washed twice with (20 mM Tris-HCl pH 7.5, 500 mM NaCl, 20 mM imidazole) and eluted with 20 mM Tris-HCl pH 7.5, 500 mM NaCl, and 500 mM imidazole. Eluted protein was dialysed overnight at 4˚C into 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM DTT and separated by size-exclusion chromatography

Thermal Denaturation assays

The stability of wild-type and mutant $TBLR1_{134-514}$ was tested in thermal denaturation assays at a final protein concentration of 2 μM and 5X SYPRO dye. The assay was performed on Biorad IQ5 ICycler using a range of 20-90˚C, increments of 1˚C and a hold time of 30 s. Melting temperatures were

obtained from the minimum of the negative first derivative of the fluorescence when plotted against temperature (-d(RFU)/dT vs T).

Crystallisation and structure refinement

The TBLR1-CTD-MeCP2 complex was formed by mixing TBLR1 (10 mg/ml) with a peptide corresponding to residues 285-309 of MeCP2 (MeCP2 NID_S; Thermo Fisher) in a protein:peptide ratio of 1:1.2. Crystals were grown by hanging drop vapor diffusion in 100 mM MOPS pH 7.5 and 18% PEG 3350 at 18ºC. Crystals were cryo-protected by supplementing buffer with 30% glycerol and flash-cooled to 100K. Diffraction data were collected at beam line I03 at Diamond Light Source. Data were indexed and integrated with XDS and merged using SCALA (36, 37). Phases were obtained by molecular replacement in PHASER (38). Coordinates of hTBLR1 (PDB ID: 4LG9, (26)) were used as a search model. The structure was built using iterative rebuilding in COOT and refinement with PHENIX using NCS restraints (39, 40).

Surface Plasmon Resonance assays

Wild-type and mutant biotin-PEG-PEG-MeCP2 peptides spanning residues 285-313 (MeCP2-NID), and biotin-PEG-PEG-MeCP2 285-309 (MeCP2-NID_S) (Thermo Fisher) were immobilized on SA sensor chips (GE Healthcare). All experiments were performed at 25˚C using a BIAcore T200 (GE Healthcare). Binding reactions were carried out in 20 mM Tris-HCl pH 7.5, 200 mM NaCl, 1 mM DTT. MeCP2 peptides (20 nM) were injected onto the chip to maximum Response Unit (RU) values of ~10. Channels for MeCP2-NID, MeCP2-NID_S and MeCP2-NID R306C peptides and a reference channel were used in parallel. TBLR1 (or a TBLR1 mutant), supplemented with 100U/ml of heparin, was injected at 30 µl/min for 30 seconds at concentrations increasing from 0.625 µM to 40 µM. Data were analysed using the BIAcore T200 Evaluation software. Binding constants were derived from steady-state analysis.

Co-localisation assays

Indicated mutations were introduced into pEGFP-C3-MeCP2 and pN1-TBL1-mCherry expression vectors, as previously (13). NIH-3T3 cells were grown on glass coverslips and transfected using JetPei (Polyplus). After 24 hours, the growth medium was changed, and after 48 hours, cells were fixed in 4% paraformaldehyde for 15 min. Cells were then washed with PBS and stained with DAPI before being mounted on slides with Prolong Diamond (Life Technologies). Finally, slides were photographed using a Zeiss Airyscan microscope.

Author Contributions

VK and MJL carried out biochemical characterisation of the complex. VK solved the crystal structure and MJL carried out cell-based assays. CT assisted with protein purification. ZT assisted with pull-down assays. VK, MJL, APB and AGC wrote the manuscript.

Acknowledgements

We thank the Xtalpod facility for crystallization screening and beam-line scientists at Diamond Light Source for support during data collection. We thank the Edinburgh Protein Production Facility for support with protein purification and biophysical characterization. We thank A. A. Jeyaprakash and Fulvia Bono for critical reading of the manuscript. AGC was supported by the Medical Research Council [G1000520/1]; Work by APB, MJL and VK was supported by the Wellcome Trust [091580,107930] and the Rett Syndrome Research Trust; The Wellcome Trust Centre for Cell Biology is supported by core funding and instrument grants [092076,095822,108504]. Crystallisation facilities were supported by a Wellcome Trust/University of Edinburgh Institutional Strategic Support Fund.

Data have been deposited in the PDB with the following accession code: 5NAF

References

1. Lewis JD*, et al.* (1992) Purification, sequence, and cellular localization of a novel chromosomal protein that binds to methylated DNA. *Cell* 69(6):905-914.

2. Skene PJ*, et al.* (2010) Neuronal MeCP2 is expressed at near histone-octamer levels and globally alters the chromatin state. *Mol Cell* 37(4):457-468.

3. Amir RE*, et al.* (1999) Rett syndrome is caused by mutations in X-linked MECP2, encoding methyl-CpG-binding protein 2. *Nat Genet* 23(2):185-188.

4. Meins M*, et al.* (2005) Submicroscopic duplication in Xq28 causes increased expression of the MECP2 gene in a boy with severe mental retardation and features of Rett syndrome. *Journal of medical genetics* 42(2):e12.

5. Van Esch H*, et al.* (2005) Duplication of the MECP2 region is a frequent cause of severe mental retardation and progressive neurological symptoms in males. *Am J Hum Genet* 77(3):442-453.

6. Ramocki MB*, et al.* (2009) Autism and other neuropsychiatric symptoms are prevalent in individuals with MeCP2 duplication syndrome. *Ann Neurol* 66(6):771-782.

7. Nan X, Meehan RR, & Bird A (1993) Dissection of the methyl-CpG binding domain from the chromosomal protein MeCP2. *Nucleic Acids Res* 21(21):4886-4892.

8. Baubec T, Ivanek R, Lienert F, & Schubeler D (2013) Methylation-dependent and -independent genomic targeting principles of the MBD protein family. *Cell* 153(2):480-492.

9. Goffin D*, et al.* (2011) Rett syndrome mutation MeCP2 T158A disrupts DNA binding, protein stability and ERP responses. *Nature neuroscience* 15(2):274-283.

10. Ho KL*, et al.* (2008) MeCP2 binding to DNA depends upon hydration at methyl-CpG. *Mol Cell* 29(4):525-531.

11. Lyst MJ & Bird A (2015) Rett syndrome: a complex disorder with simple roots. *Nature reviews. Genetics* 16(5):261-275.

12. Kokura K*, et al.* (2001) The Ski protein family is required for MeCP2-mediated transcriptional repression. *J Biol Chem* 276(36):34115-34121.

13. Lyst MJ*, et al.* (2013) Rett syndrome mutations abolish the interaction of MeCP2 with the NCoR/SMRT co-repressor. *Nature neuroscience* 16(7):898-902.

14. Stancheva I, Collins AL, Van den Veyver IB, Zoghbi H, & Meehan RR (2003) A mutant form of MeCP2 protein associated with human Rett syndrome cannot be displaced from methylated DNA by notch in Xenopus embryos. *Mol Cell* 12(2):425-435.

15. Brown K*, et al.* (2016) The molecular basis of variable phenotypic severity among common missense mutations causing Rett syndrome. *Hum Mol Genet* 25(3):558-570.

16. Heckman LD, Chahrour MH, & Zoghbi HY (2014) Rett-causing mutations reveal two domains critical for MeCP2 function and for toxicity in MECP2 duplication syndrome mice. *Elife* 3.

17. Collins AL*, et al.* (2004) Mild overexpression of MeCP2 causes a progressive neurological disorder in mice. *Hum Mol Genet* 13(21):2679-2689.

18. Guenther MG, Barak O, & Lazar MA (2001) The SMRT and N-CoR corepressors are activating cofactors for histone deacetylase 3. *Mol Cell Biol* 21(18):6091-6101.

19. Li J*, et al.* (2000) Both corepressor proteins SMRT and N-CoR exist in large protein complexes containing HDAC3. *EMBO J* 19(16):4342-4350.

20. Yoon HG*, et al.* (2003) Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1. *EMBO J* 22(6):1336-1346.

21. Zhang J, Kalkum M, Chait BT, & Roeder RG (2002) The N-CoR-HDAC3 nuclear receptor corepressor complex inhibits the JNK pathway through the integral subunit GPS2. *Mol Cell* 9(3):611-623.

22. Oberoi J*, et al.* (2011) Structural basis for the assembly of the SMRT/NCoR core transcriptional repression machinery. *Nature structural & molecular biology* 18(2):177-184.

23. Ebert DH*, et al.* (2013) Activity-dependent phosphorylation of MeCP2 threonine 308 regulates interaction with NCoR. *Nature* 499(7458):341-345.

24. Shahbazian M*, et al.* (2002) Mice with truncated MeCP2 recapitulate many Rett syndrome features and display hyperacetylation of histone H3. *Neuron* 35(2):243-254.

25. Wu XH, Wang Y, Zhuo Z, Jiang F, & Wu YD (2012) Identifying the hotspots on the top faces of WD40-repeat proteins from their primary sequences by beta-bulges and DHSW tetrads. *PloS one* 7(8):e43005.

26. Xu C*, et al.* (2013) Crystal structure of TBL1XR1 WD40 repeats. *PDB ID: 4LG9*.

27. Chen VB*, et al.* (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D* 66:12-21.

28. Stirnimann CU, Petsalaki E, Russell RB, & Muller CW (2010) WD40 proteins propel cellular networks. *Trends Biochem Sci* 35(10):565-574.

29. Landau M*, et al.* (2005) ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* 33(Web Server issue):W299-302.

30. Migliori V*, et al.* (2012) Symmetric dimethylation of H3R2 is a newly identified histone mark that supports euchromatin maintenance. *Nature structural & molecular biology* 19(2):136-144.

31. Xu C*, et al.* (2010) Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). *Proc Natl Acad Sci U S A* 107(45):19266-19271.

32. Christodoulou J, Grimm A, Maher T, & Bennetts B (2003) RettBASE: The IRSA MECP2 variation database-a new mutation database in evolution. *Hum Mutat* 21(5):466-472.

33. Firth HV*, et al.* (2009) DECIPHER: Database of Chromosomal Imbalance and Phenotype in Humans Using Ensembl Resources. *Am J Hum Genet* 84(4):524-533.

34. O'Roak BJ*, et al.* (2012) Multiplex targeted sequencing identifies recurrently mutated genes in autism spectrum disorders. *Science* 338(6114):1619-1622.

35. Laskowski RA*, et al.* (2016) Integrating population variation and protein structural analysis to improve clinical interpretation of missense variation: application to the WD40 domain. *Hum Mol Genet* 25(5):927-935.

36. Evans P (2006) Scaling and assessment of data quality. *Acta Crystallogr D Biol Crystallogr* 62(Pt 1):72-82.

37. Kabsch W (2010) Integration, scaling, space-group assignment and post-refinement. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):133-144.

38. McCoy AJ*, et al.* (2007) Phaser crystallographic software. *J Appl Crystallogr* 40(Pt 4):658-674.

39. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213-221.

40. Emsley P & Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126-2132.

Table 1. Binding constants derived from non-linear regression analysis of SPR data; s.e., standard error; n.b., no binding. R^2, weighted residual of fit.

	MeCP2-NID				MeCP2-NID$_S$			
TBLR1	**K_D ± s.e. (μM)**	**RU max**	**R^2**	**χ2**	**K_D ± s.e. (μM)**	**RU max**	**R^2**	**χ2**
Structure-based mutants								
Wild-type	9.5 ± 0.5	81.5	0.998	0.80	12.9 ± 0.8	105.7	0.998	1.41
E171A	≥ 40	32.5	0.873	5.09	≥ 40	77.5	0.995	0.82
E171Q	≥ 40	60.3	0.989	1.20	≥ 40	72.4	0.996	0.74
C214S	6.6 ± 0.4	97.45	0.998	1.83	8.9 ± 0.5	134.4	0.998	2.74
D313N	n.b.	-	-	-	n.b.	-	-	-
E351A	32 ± 5	57.2	0.993	0.88	≥ 40	83.0	0.998	0.28
E351D	11.4 ± 0.5	84.3	0.999	0.53	16.3 ± 0.7	114.5	0.999	0.69
D369A	n.b.	-	-	-	n.b.	-	-	-
Y446F	10.1 ± 0.5	96.9	0.999	1.01	13.0 ± 0.7	128.2	0.999	1.55
DECIPHER mutants								
D369E	29 ± 3.1	18.9	0.996	0.07	≥ 40	27.6	0.985	0.29
P444R	24 ± 2.3	46.1	0.996	0.45	31.6 ± 4	69.1	0.994	1.17

Figure Legends

Figure 1. TBL1/TBLR1 and MeCP2 interact with a common fragment of NCoR1.
(**a**) Domain overview of MeCP2 (orange) and the NCoR/SMRT complex components, TBLR1 (blue) and NCoR1 (brown). Above is a sequence of MeCP2 NID (orange) with residues mutated in RTT highlighted in yellow. Above the sequence are RTT mutations used in this study. The T308 phosphorylation site is marked with a P in a red circle. A previously characterized interaction between TBL1 and NCoR1 is indicated with gray dotted lines. Domains are annotated as: MBD, methylated DNA binding domain; NID, NCoR interaction domain; WD40, domain containing 8 WD40 motif repeats; DAD, deacetylase activating domain. Brown lines indicate NCoR1 deletion mutants. Results of binding interactions between NCoR1 fragments with TBL1 and MeCP2 are summarized; n.d., not determined. (**b**) Pull-down of FLAG-NCoR1 fragments from HeLa cell extracts by MeCP2-NID peptide. (**c**) Co-IP of endogenous TBLR1 with FLAG-NCoR1 fragments from transfected HeLa cells.

Figure 2. MeCP2 binds directly to TBL1 C-terminal domain.
(**a**) Overview of TBL1/ TBLR1 domain structure (blue). Blue lines show deletion mutants of TBL1. Results from pull-down assays are summarized beside the overview of deletion mutants. (**b**) Pull-down assay of full-length and deletion fragments of TBL1 from HeLa cell extracts with MeCP2 NID peptide. B indicates samples bound to beads containing biotinylated peptide. Peptides were run as a control in adjacent lanes (P). (**c**) Pull-down assay of purified TBLR1-CTD with wild-type MeCP2 NID peptide and peptides encoding RTT mutations or phosphoThr308 modification. (**d**) Binding interaction between MeCP2 peptides and TBLR1-CTD by surface plasmon resonance. Steady state binding curves were derived from binding reactions with MeCP2 NID (blue), MeCP2 NID_S (gray) and MeCP2 NID R306C mutant (red) peptides.

Figure 3. TBL1 chimeras allow identification of MeCP2 binding residues.
(**a**) Overview of mouse-sponge chimeras. Wild-type and mutant MeCP2-NID peptides were used to pull-down FLAG-tagged TBL1 chimeras. Binding interactions with MeCP2 are summarized next to the schema. (**b**) Pull-down assays of wild-type and mutant MeCP2 NID with FLAG-tagged mouse TBL1 mutants with individual amino acids converted to residues found in sponge.

Figure 4. Co-crystal structure of the TBLR1-MeCP2 complex.
(**a**) Cartoon representation of TBLR1-CTD (blue) with MeCP2 NID (orange) shown as ribbon. WD40 blades are coloured shades of blue, with a gradient of dark to light from N- to C-terminus. (**b**) Surface representation of TBLR1-CTD with MeCP2 NID. The view is rotated 90° around the horizontal axis compared to (**a**) and WD40 repeats 1-3 are removed for clarity. (**c**) Surface representation of TBLR1 with the view as in (a). Yellow indicates the area buried by MeCP2. (**d**) Electrostatic surface representation of TBLR1-CTD (APBS) colored red to blue from -5.5 to +5.5 kT/e. (**e**) Surface conservation analysis of TBLR1 (CONSURF), colored from blue (variable) to purple (conserved).

Figure 5. RTT mutations disrupt key interactions between MeCP2 and TBLR1.
(**a**) Close-up of MeCP2 NID (orange) bound to TBLR1 (blue). Only residue side chains that make contacts with MeCP2 are shown. Hydrogen bonds are grey dotted lines. (**b**) The same view, rotated 180° around the vertical axis. (**c-f**) MeCP2 is shown with RTT mutations modeled and wild-type side chains shown as transparent ghosts: (**c**) R306C (**d**) K304E, (**e**) P302R and (**f**) K305R. Orange arrows indicate the mutated residue.

Figure 6. TBL1/TBLR1 surface mutations do not bind MeCP2.
(**a**) Map of mutations in TBLR1 and TBL1. (**b**) Steady state SPR binding curves of wild-type (blue) and mutant (gray and black) TBLR1 proteins interacting with immobilized MeCP2 NID peptide. Wild-type TBLR1 with R306C NID peptide (red) is a negative control. (**c**) Recruitment of TBL1-mCherry to heterochromatin foci in NIH-3T3 cells by EGFP-MeCP2. DAPI staining (blue), green and red fluorescence and a merged image are shown for wild-type and mutant proteins. The scale bar represents 5 micron. Wild-type cells show colocalistion in 63% of cells (n=27). Mutants E184A, E364A, E382A and D326N show colocalisation in 0-1 cells in 14. Y459F shows colocalisation in 55% of cells (n=27).

Figure 7. Missense mutations associated with developmental delay block MeCP2 binding to TBL1/TBLR1.
(**a**) Map of mutations in TBLR1 and TBL1. (**b**) Pull-down assays of equivalent mutations in TBL1 with MeCP2 NID peptide or a mutant peptide containing the K305R RTT mutation. (**c**) Recruitment of TBL1-mCherry (red) and developmental delay mutants by EGFP-MeCP2 (green) to heterochromatin foci (blue) in NIH-3T3 cells. The scale bar represents 5 micron.

Figure 1

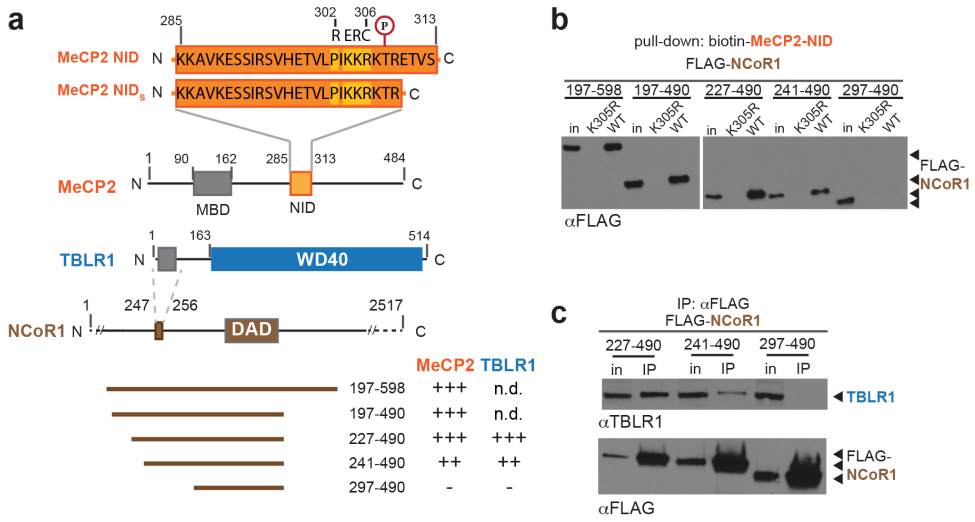

a
285
302
306
313
R ERC
MeCP2 NID N KKAVKESSIRSVHETVLPIKKRKTRETVS C
MeCP2 NID$_s$ N KKAVKESSIRSVHETVLPIKKRKTR C
MeCP2
1
90
162
285
313
484
N
C
MBD
NID
TBLR1
1
163
514
WD40
NCoR1
1
247
256
2517
DAD
MeCP2 TBLR1
197-598 +++ n.d.
197-490 +++ n.d.
227-490 +++ +++
241-490 ++ ++
297-490 - -
b
pull-down: biotin-MeCP2-NID
FLAG-NCoR1
197-598
197-490
227-490
241-490
297-490
in
K305R
WT
αFLAG
FLAG-
NCoR1
c
IP: αFLAG
FLAG-NCoR1
227-490
241-490
297-490
in
IP
TBLR1
αTBLR1
FLAG-
NCoR1
αFLAG

Figure 2

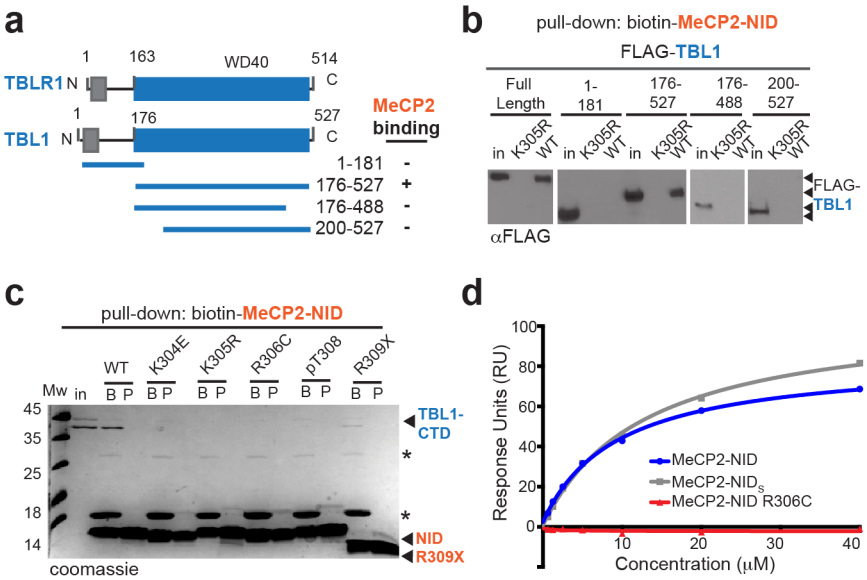

a
TBLR1
TBL1
WD40
MeCP2 binding
1-181 -
176-527 +
176-488 -
200-527 -
b
pull-down: biotin-MeCP2-NID
FLAG-TBL1
Full Length
1-181
176-527
176-488
200-527
αFLAG
FLAG-TBL1
c
pull-down: biotin-MeCP2-NID
WT
K304E
K305R
R306C
pT308
R309X
coomassie
TBL1-CTD
NID R309X
d
Response Units (RU)
Concentration (μM)
MeCP2-NID
MeCP2-NIDS
MeCP2-NID R306C

Figure 3

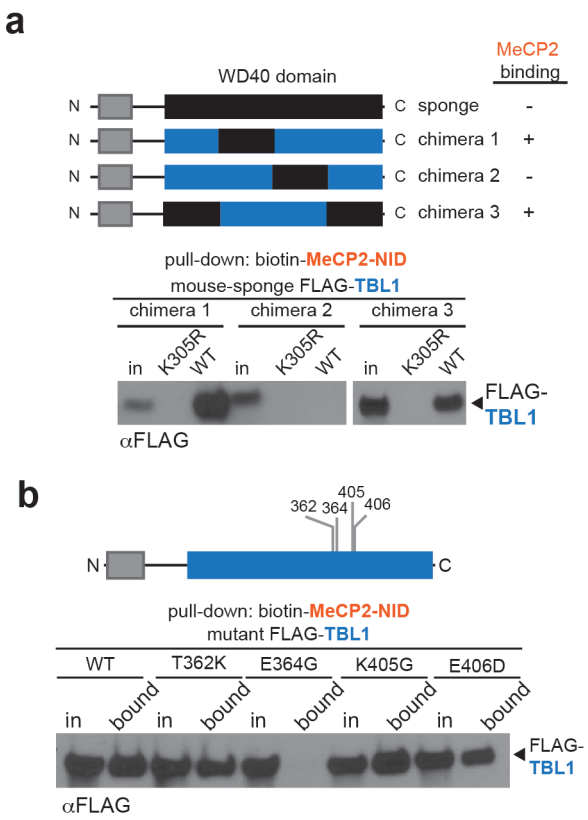

a
MeCP2 binding
WD40 domain
N
C
sponge
-
chimera 1
+
chimera 2
-
chimera 3
+
pull-down: biotin-MeCP2-NID
mouse-sponge FLAG-TBL1
chimera 1
chimera 2
chimera 3
in
K305R
WT
FLAG-TBL1
αFLAG
b
405
362 364
406
pull-down: biotin-MeCP2-NID
mutant FLAG-TBL1
WT
T362K
E364G
K405G
E406D
in
bound
FLAG-TBL1
αFLAG

Figure 4

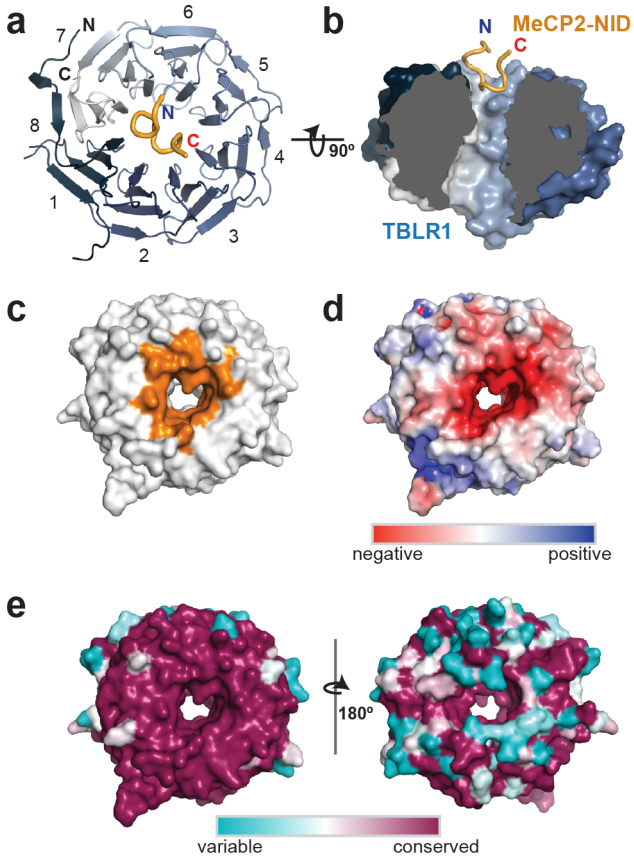

a
b
c
d
e
N
C
6
7
5
8
4
1
2
3
90°
MeCP2-NID
TBLR1
negative
positive
180°
variable
conserved

Figure 5

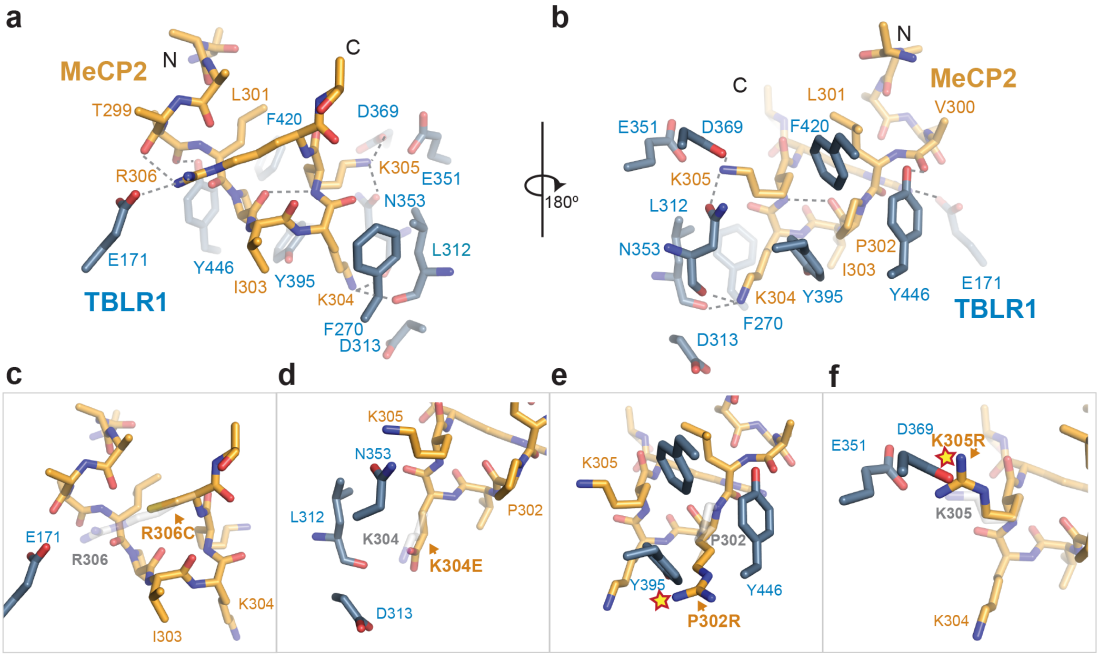

a
MeCP2
N
C
L301
T299
F420
D369
R306
K305
E351
N353
E171
Y446
I303
Y395
K304
L312
F270
D313
TBLR1
180°
b
N
MeCP2
C
L301
V300
E351
D369
F420
K305
L312
N353
P302
I303
K304
Y395
Y446
E171
F270
D313
TBLR1
c
E171
R306C
R306
K304
I303
d
K305
N353
L312
P302
K304
K304E
D313
e
K305
P302
Y395
Y446
P302R
f
E351
D369
K305R
K305
K304

Figure 6

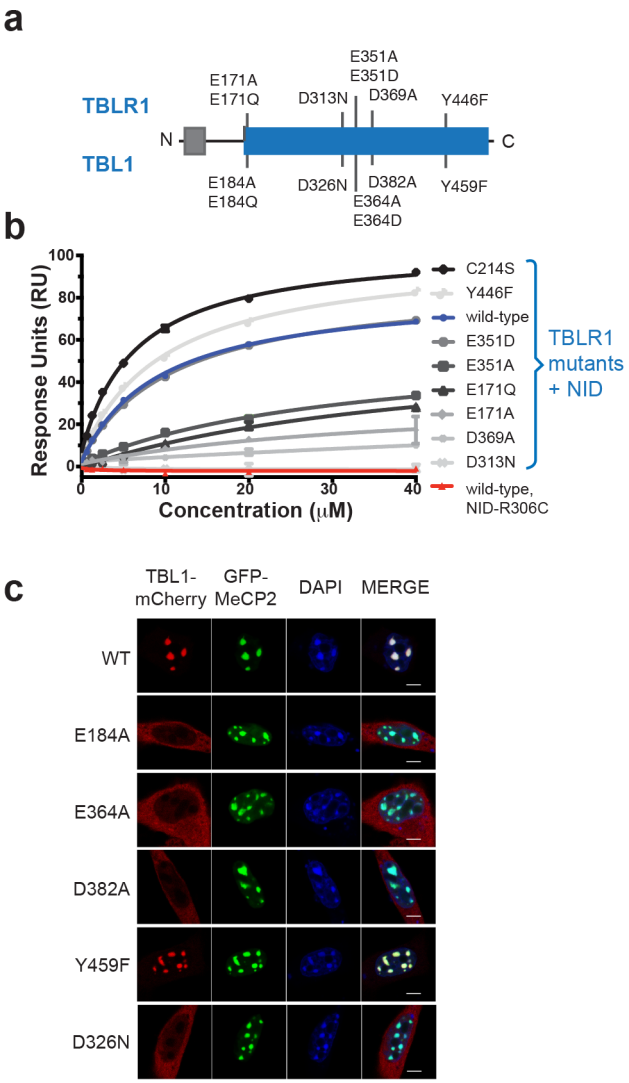
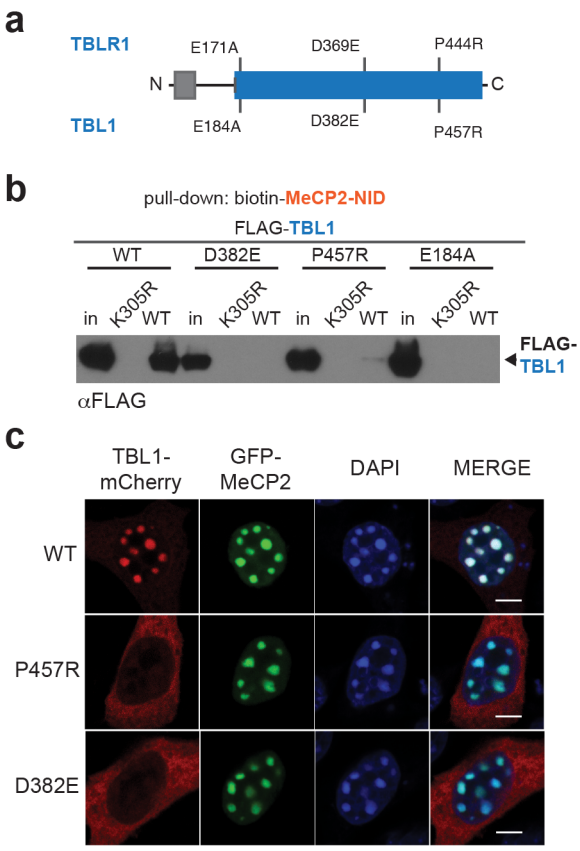

a
TBLR1
TBL1
N
C
E171A
E171Q
E184A
E184Q
D313N
D326N
E351A
E351D
D369A
D382A
E364A
E364D
Y446F
Y459F
b
Response Units (RU)
Concentration (μM)
C214S
Y446F
wild-type
E351D
E351A
E171Q
E171A
D369A
D313N
wild-type,
NID-R306C
TBLR1
mutants
+ NID
c
TBL1-
mCherry
GFP-
MeCP2
DAPI
MERGE
WT
E184A
E364A
D382A
Y459F
D326N

Figure 7


a
TBLR1
E171A
D369E
P444R
N
C
TBL1
E184A
D382E
P457R
b
pull-down: biotin-MeCP2-NID
FLAG-TBL1
WT
D382E
P457R
E184A
in
K305R
WT
FLAG-TBL1
αFLAG
c
TBL1-mCherry
GFP-MeCP2
DAPI
MERGE
WT
P457R
D382E

Supplemental Information

Supplementary Methods

Fluorescence anisotropy assays

All assays were carried out in a final volume of 100 µl in black opaque 96-well plates (Greiner Bio One) using a SpectraMax M5 multimode plate reader at room temperature. Fluorescein-labeled MeCP2 298-309 peptide, at a final concentration of 100 nM, was titrated against protein concentrations ranging from 50 nM to 40 µM. Data were analyzed using non-linear regression with eq. 1 where R_f and R_b are the anisotropies of the free and bound receptor, respectively, L_T is the total ligand concentration, and K_D is the dissociation constant.

$$\boldsymbol{Y = R_f + \left(R_b - R_f\right) * \frac{\left(L_T + K_D + X - \sqrt{(-L_T - K_D - X)^2 - 4 * K_D * X}\right)}{2 * L_T}} \quad \textbf{(1)}$$

Table S1. Data collection and refinement statistics

	TBLR1-CTD-MeCP2-NID$_S$
Data Collection	
Beamline	DLS I03
Wavelength (Å)	0.976250
Resolution range Å (high resolution shell)*	49.6 - 2.5 (2.6 - 2.5)
Space group	*P* 1
Unit cell a,b,c (Å)	42.6 58.6 155.1
α, β, γ (°)	97.6 91.2 90.2
Total reflections	120672 (10968)
Unique reflections	50648 (4583)
Multiplicity	2.4 (2.4)
Completeness (%)	98 (88)
Mean I/σ(I)	9.9 (2.3)
Wilson B-factor (Å^2)	33.5
R-merge	0.08 (0.4)
CC1/2	0.97 (0.69)
Refinement	
Reflections used in refinement	50632 (4583)
R_{work}	0.20
R_{free}	0.25
Number of non-hydrogen atoms	10400
protein	9619
peptide	153
solvent	619
r.m.s. (bonds)	0.004
r.m.s (angles)	0.72
Ramachandran favored (%)	95.3
Ramachandran allowed (%)	4.3
Ramachandran outliers (%)	0.4
Clashscore	7.1
Average B-factor (Å^2)	50.2
protein	52.6
peptide	46.1
solvent	52.3

*Statistics for the highest-resolution shell are shown in parentheses.

Supplemental Figures

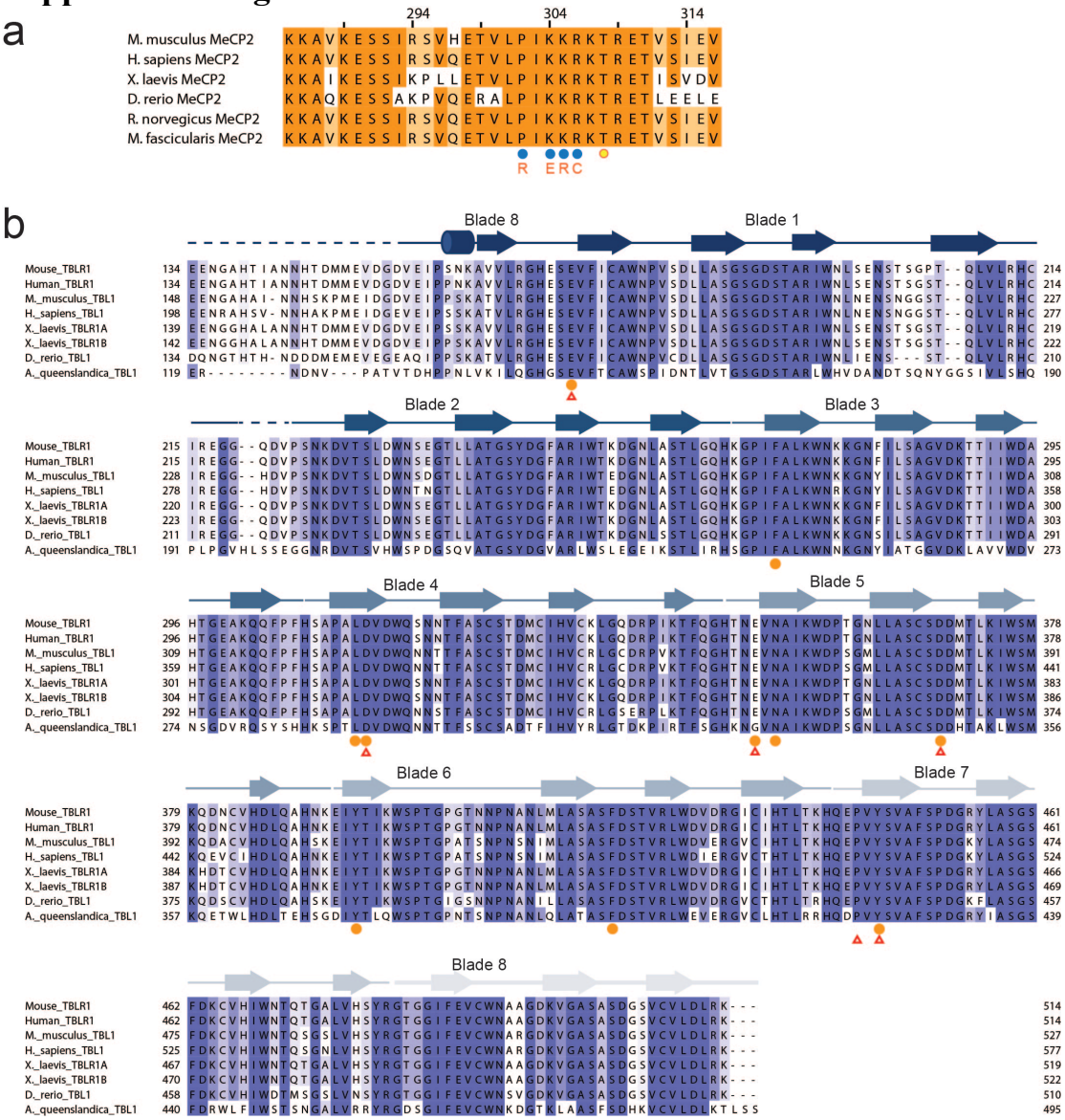


Figure S1

(**a**) Sequence alignment for MeCP2 NID domain. Residues that interact with TBLR1 are marked with blue dots. T308, which can be phosphorylated is marked with a yellow dot. Known RTT missense mutations used in this study are marked below. (**b**) Sequence alignment of TBL1 and TBLR1 from vertebrates and sponge. Secondary structure elements are shown above the sequence. Residues interacting with MeCP2 are marked with orange dots. Residues mutated in this study are marked with red triangles.

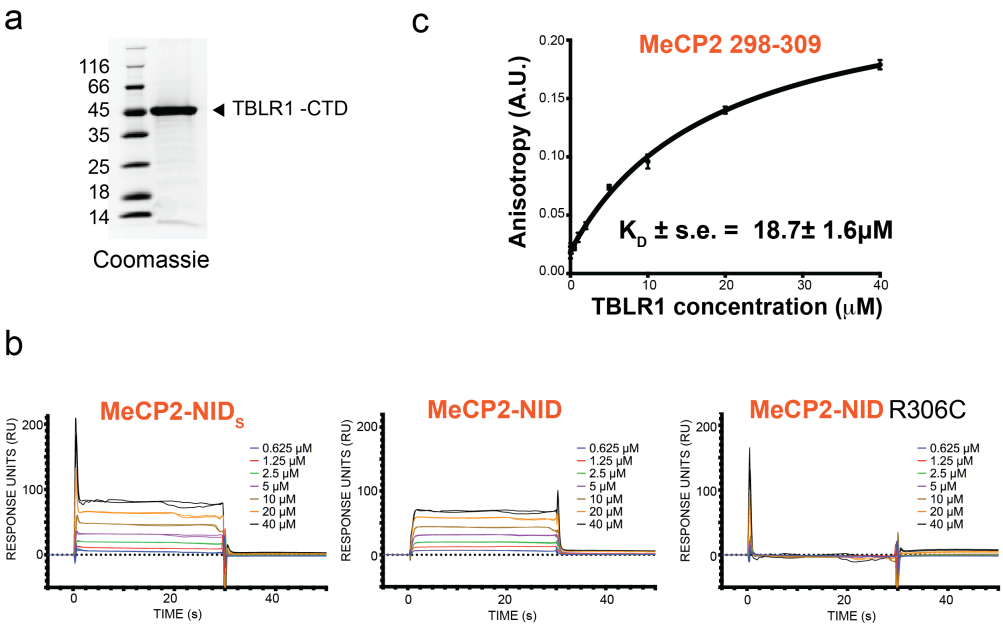


Figure S2

(**a**) Coomassie stained gel of purified TBLR1-CTD. (**b**) Raw surface plasmon resonance traces for wild-type TBLR1-CTD with MeCP2 mutant and wild-type peptides. The key indicates the concentrations of TBLR1 used in a dilution series. Binding curves in Fig. 2 were calculated from equilibrium measurements from these data. (**c**) Fluorescence anisotropy binding with minimal 11mer peptide of MeCP2 (residues 298-309). Anisotropy is measured in arbitrary units (A.U.). The binding constant was derived from fitting three independent experiments and is expressed as the mean K_D ± standard error (s.e.).

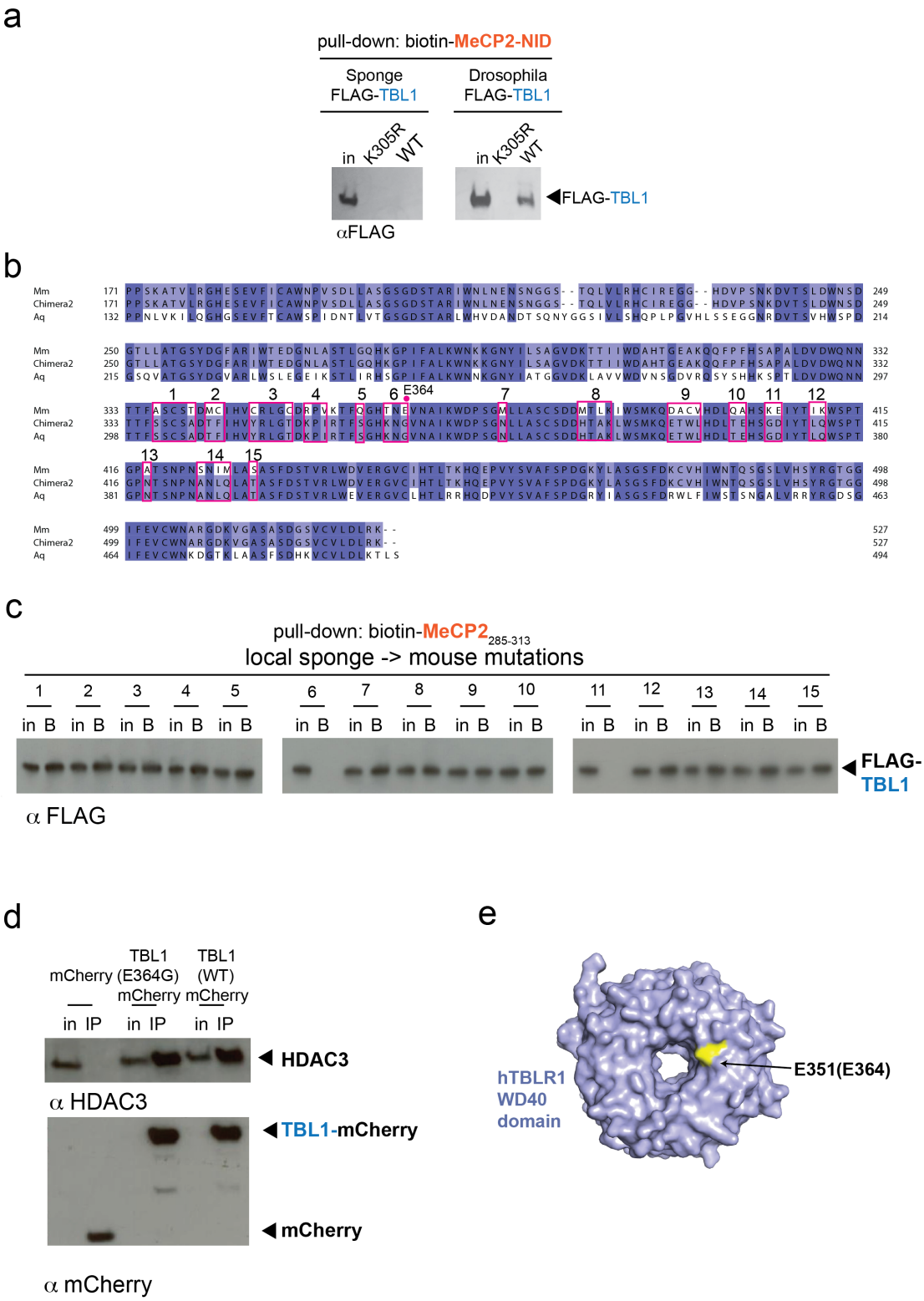


Figure S3

(**a**) Pull-down assay of biotinylated-MeCP2-NID wild-type or mutant peptide with TBL1 from sponge and *Drosophila*. (**b**) Sequence alignment of TBL1 from mouse and sponge, as well as TBL1-chimera2, indicating segments of the sequence that were used for mutagenesis analysis. (**c**) Pull-down assays between MeCP2 NID and local TBL1 mouse-> sponge mutants from HeLa cell extracts. (**d**) Co-IP of HDAC3 with wild-type and mutant TBL1. (**e**) Location of residue equivalent to E364 on the human TBLR1 structure (PDBid 4LG9).

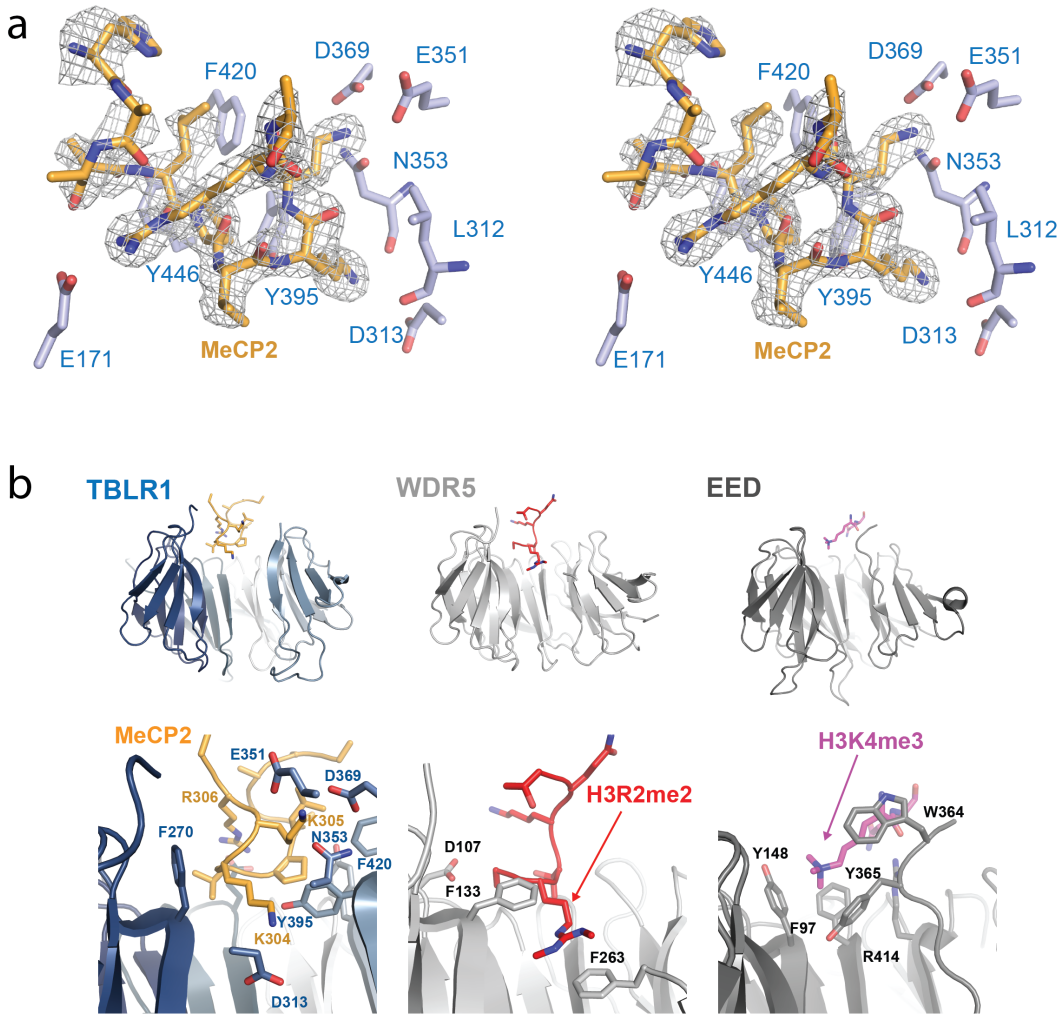


Figure S4
(**a**) Divergent stereo view of the MeCP2 NID peptide (yellow) bound to TBLR1 (blue) with Fo-Fc omit map calculated from coordinates where the MeCP2 peptide was omitted. This map is contoured at 2.5σ. (**b**) A comparison of peptide recognition in two histone H3 binding WD40 domains, WDR5 (PDBid 4A7J) and EED (PDBid 3K26) that feature recognition of a basic motif. The orientation of the viewpoint is rotated 90° around the y-axis with respect to (**a**). Each WD40 domain was superposed using the SSM function in COOT and are aligned primarily on the first 4 repeats. Root mean square deviations of fit are 2.6 Å and 2.7 Å respectively. WD40 repeats 4 and 5 have been removed for clarity. Methylated arginine and lysine residues from histone H3 are recognized by aromatic cages, in contrast to the ionic interactions observed with basic residues in MeCP2.

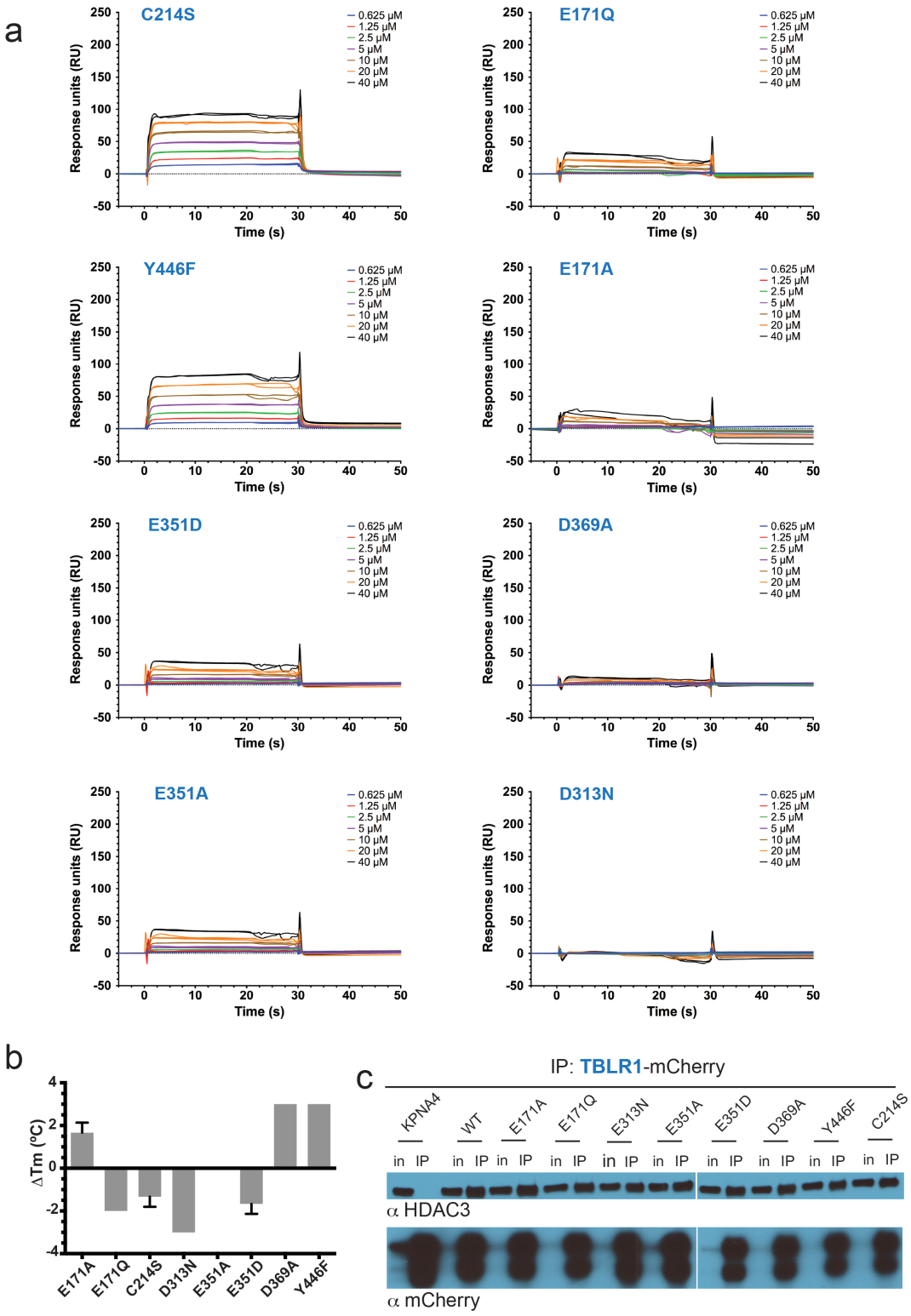


Figure S5

(**a**) Raw surface plasmon resonance traces for TBLR1-CTD mutants with MeCP2 NID peptide. (**b**) Alterations in melting temperature (Tm) of TBLR1 mutants using a thermal denaturation assay. Change in Tm is compared to wild-type. (**c**) co-IP of mCherry TBLR1 or mCherry importin α4 (KPNA4) with NCoR/SMRT components. Samples tested were: KPNA4 (negative control, has a similar molecular mass as TBLR1), wild-type TBLR1 (positive control) and TBLR1 point mutants. Co-purification of NCoR/SMRT was determined by probing with α-HDAC3; in = input; IP = immunoprecipitate.

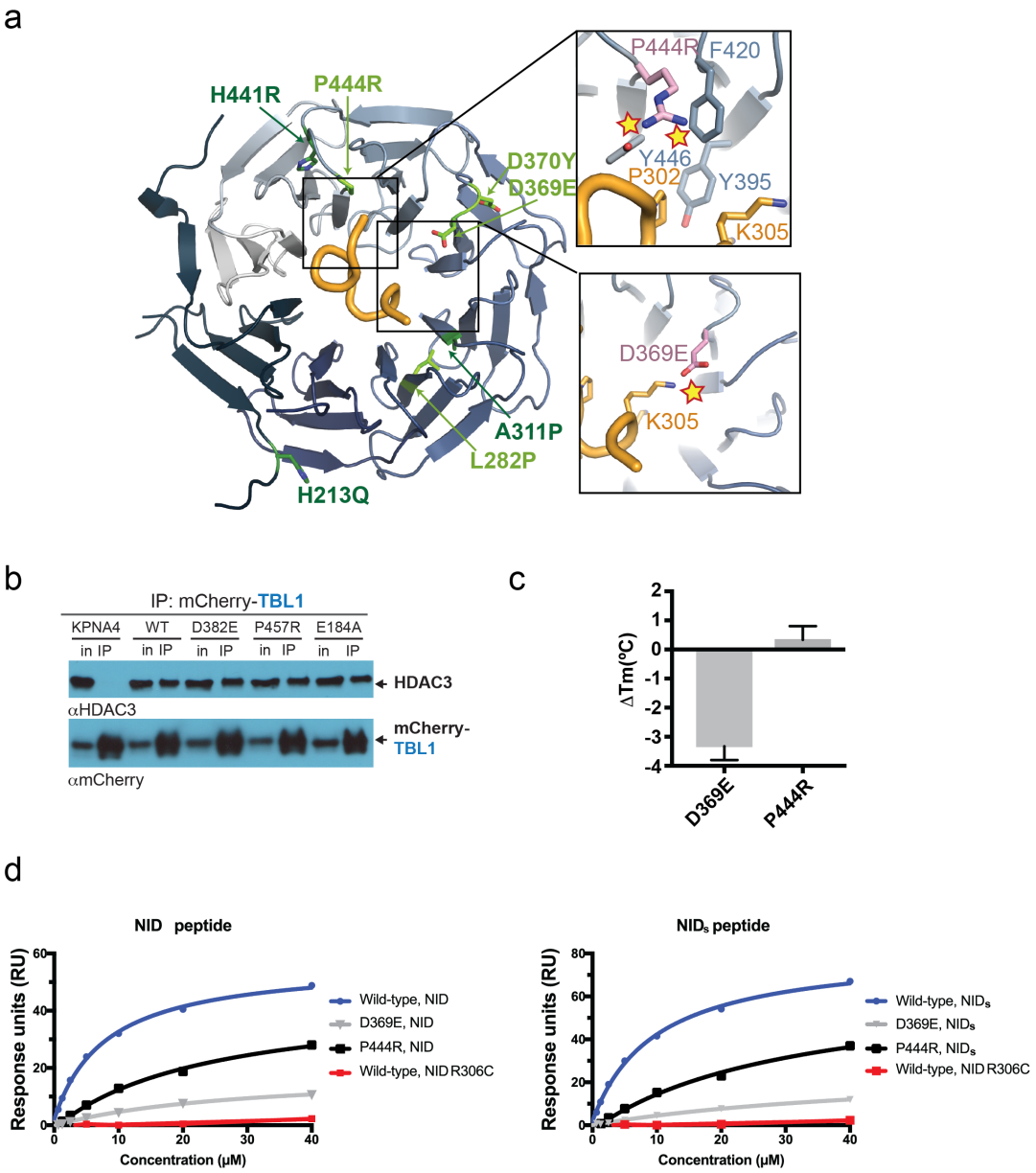


Figure S6

(**a**) Overview of TBLR1 mutations associated with developmental delay without (dark green) or with (light green) autism. Models of P444R and D369E mutations (pink) are shown in inset boxes. (**b**) Co-IP of mCherry TBLR1 or mCherry importin α4 (KPNA4) with NCoR/SMRT components. Samples tested were: KPNA4 (negative control), wild-type TBL1 (positive control) and TBL1 point mutants. Residues in TBLR1 that are equivalent to these mutations are D369E, P444R and E171A. Co-purification of NCoR/SMRT was determined by probing with α-HDAC3; in = input; IP = immunoprecipitate. (**c**) Thermal denaturation assays for developmental delay mutants showing change in Tm with respect to wild-type. (**d**) SPR binding curves for NID and NID_S peptides with developmental delay mutants. Curves for wild-type TBLR1 protein are shown with wild-type (blue) and RTT mutant peptides (red). Mutant TBLR1 proteins with wild-type peptide are shown in gray and black.